Polymorphisms in the *hMSH2* Gene and the Risk of Primary Lung Cancer

Chi Young Jung,[1] Jin Eun Choi,[2] Jung Min Park,[2] Myung Hwa Chae,[2] Hyo-Gyoung Kang,[2] Kyung Mee Kim,[2] Su Jeong Lee,[2] Won Kee Lee,[3] Sin Kam,[3] Seung Ick Cha,[1] Chang Ho Kim,[1] Sung Beom Han,[4] Tae Hoon Jung,[1] Su Han Jeon,[5] and Jae Yong Park[1,2]

[1]Department of Internal Medicine, [2]Cancer Research Institute, and [3]Department of Preventive Medicine, Kyungpook National University; [4]Department of Internal Medicine, School of Medicine, Keimyung University; and [5]Department of General Surgery, Kyungpook National University Hospital, Daegu, Korea

Abstract

Polymorphisms in the DNA repair genes may be associated with differences in the capacity to repair DNA damage, and so this can influence an individual's susceptibility to lung cancer. To test this hypothesis, we investigated the association of *hMSH2* −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12−6T>C genotypes and their haplotypes with the risk of lung cancer in a Korean population. The *hMSH2* genotypes were determined in 432 lung cancer patients and in 432 healthy controls who were frequency matched for age and gender. The *hMSH2* haplotypes were estimated based on a Bayesian algorithm using the Phase program. The presence of at least one IVS10+12G allele was associated with a significantly decreased risk of adenocarcinoma, as compared with the IVS10+12AA genotype [adjusted odds ratio (OR), 0.59; 95% confidence interval (95% CI), 0.40-0.88; $P = 0.01$], and the presence of at least one IVS12-6C allele was associated with a significantly increased risk of adenocarcinoma, as compared with the IVS12-6TT genotype (adjusted OR, 1.52; 95% CI, 1.02-2.27; $P = 0.04$). Consistent with the results of the genotyping analysis, the TGGT haplotype with no risk allele was associated with a significantly decreased risk of adenocarcinoma, as compared with the TCAC haplotype with two risk allele [i.e., IVS10+12A and IVS12-6C allele; adjusted OR, 0.49; 95% CI, 0.30-0.78; $P = 0.003$ and P_c (Bonferroni corrected P value) = 0.012]. The effect of the *hMSH2* haplotypes on the risk of adenocarcinoma was statistically significant in the never smokers and younger individuals (adjusted OR, 0.45; 95% CI, 0.27-0.75; $P = 0.002$ and $P_c = 0.004$; and adjusted OR, 0.44; 95% CI, 0.23-0.85; $P = 0.014$ and $P_c = 0.028$, respectively) but not in the ever-smokers and older individuals. These results suggest that the *hMSH2* polymorphisms and their haplotypes may be an important genetic determinant of adenocarcinoma of the lung, particularly in never smokers. (Cancer Epidemiol Biomarkers Prev 2006;15(4):762–8)

Introduction

Although cigarette smoking is the major cause of lung cancer, only a small fraction of smokers develop the disease, and this suggests that genetic factors contribute to the risk of lung cancer. This genetic susceptibility may result from inherited polymorphisms in the genes involved in the carcinogen metabolism and in the repair of DNA damage (1, 2).

DNA repair systems are of fundamental importance for the maintenance of genomic integrity in the face of replication errors, environmental carcinogens, and the cumulative effects of age, and their inactivation can dramatically increase the susceptibility to cancer (3, 4). In humans, >70 genes are involved in the five major DNA repair pathways: nucleotide excision repair, base excision repair, mismatch repair, homologous recombinational repair, and nonhomologous end joining (3, 4).

Molecular epidemiologic studies have shown considerable interindividual variation in the DNA repair capacity in the general population. Individuals with a suboptimal DNA repair capacity are at an increased risk of cancers such as lung cancer and squamous cell carcinoma of the head and neck (5, 6). Polymorphisms in the DNA repair genes may contribute to the DNA repair capacity variations in the general population. Therefore, it has been hypothesized that inherited polymorphisms in the DNA repair genes may modulate the susceptibility to lung cancer. To test this hypothesis, we have previously studied the contribution of polymorphisms in the DNA repair genes to the risk of lung cancer in a Korean population (7-9).

A highly conserved set of mismatch repair proteins is primarily responsible for the correction of replication errors (base-base or insertion-deletion mismatches) that are caused by DNA polymerase errors (10, 11). Genetic and epigenetic inactivation of the mismatch repair genes has been implicated in the etiology of hereditary nonpolyposis colorectal cancer syndrome and also in a wide variety of sporadic tumors such as colorectal, ovarian, and endometrial cancers (12, 13). However, the pathogenic role of the mismatch repair genes in the environment-induced cancers such as lung cancer has not been well defined (14, 15). In addition to their established role in the repair of postreplicative DNA errors, mismatch repair proteins are also involved in a variety of other vital cellular processes such as the induction of apoptosis in response to exogenous DNA damage (16, 17) and the transcription-coupled nucleotide excision repair of bulky DNA adducts (18, 19). Therefore, a subtle defect in the DNA repair capacity that is caused by functional polymorphisms in the mismatch repair genes that are neither necessary nor sufficient for the development of lung cancer could place some individuals at an increased risk of lung cancer.

The *hMSH2* gene is one of the mismatch repair genes and it encodes the human homologue of the bacterial MutS protein, which is responsible for recognizing DNA mismatches (10, 11).

Received 10/26/05; revised 2/1/06; accepted 2/14/06.

Grant support: The Regional Technology Innovation Program of The Ministry of Commerce, Industry, and Energy grant RTI04-01-01.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: C.Y. Jung and J.E. Choi contributed equally to this work.

Requests for reprints: Jae Yong Park, Department of Internal Medicine, Kyungpook National University Hospital, Samduk 2Ga 50, Daegu 700-412, Korea. Phone: 82-53-420-5536; Fax: 82-53-426-2046. E-mail: jaeyong@kyungpook.ac.kr

Copyright © 2006 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-05-0834

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

To date, several polymorphisms in the *hMSH2* gene have been reported (refs. 20-33; Table 1). Whereas the functional effects of these polymorphisms are not known, we hypothesized that some of these variants, and particularly their haplotypes, may have an effect on the DNA repair capacity, and this can modulate the susceptibility to lung cancer. To test this hypothesis, we conducted a case-control study to evaluate the association between the *hMSH2* genotypes/haplotypes and the risk of lung cancer. Among the identified polymorphisms in the *hMSH2* gene, we evaluated the association of the −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12-6T>C polymorphisms with lung cancer because the other polymorphisms were not detected in a preliminary study that consisted of 27 lung cancer cases and 27 healthy controls.

Materials and Methods

Study Population. This case-control study included 432 lung cancer patients and 432 healthy controls. The details of the study population have been described elsewhere (34, 35). In brief, the eligible cases included all the patients who were newly diagnosed with primary lung cancer at Kyungpook National University Hospital, Daegu, Korea from January 2001 to February 2002. There were no age, gender, histologic, or stage restrictions but those patients who had a prior history of cancers were excluded from this study. The cases included 210 (48.6%) squamous cell carcinomas, 141 (32.6%) adenocarcinomas, 73 (16.9%) small-cell carcinomas, and 8 (1.9%) large-cell carcinomas. The control subjects were randomly selected from a pool of healthy volunteers who had visited the general health check-up center at Kyungpook National University Hospital during the same period. The control subjects were frequency matched (1:1) to the cancer cases on the basis of gender and age (±5 years). All the cases and the controls were ethnic Koreans and they resided in Daegu City or in the surrounding regions. A detailed questionnaire was completed for each patient and each control by a trained interviewer. The questionnaire included information on the average number of cigarettes they smoked daily and the number of years the subjects had been smoking. For the smoking status of the subjects, a person who had smoked at least once a day for >1 year during his or her lifetime was regarded as a smoker. A former smoker was defined as a person who had stopped smoking at least 1 year before the diagnosis of lung cancer in the case of the patients and 1 year before the date signed on an informed consent form for the blood sample collection in the case of the controls. The cumulative cigarette dose (pack-years) was calculated by using the following formula: pack-years = (packs per day) × (years smoked).

Genotyping. Genomic DNA was extracted from peripheral blood lymphocytes by proteinase K digestion and phenol/chloroform extraction. The *hMSH2* −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12-6T>C genotypes were determined by using a PCR-restriction fragment length polymorphism assay. The PCR primers were designed based on the GenBank reference sequence (accession no. AY601851). The PCR primers for the −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12-6T>C polymorphisms were 5′-GAAACGCAGCCCTGGAAGCTA(G→A)A-3′ (forward) and 5′-AAACCTCCTCACCTCCTG-GTTG-3′ (reverse); 5′-GACCGGGGCGACTTCTATAC-3′ (forward) and 5′-AAAGGAGCCGCGCCACAAGG-3′ (reverse); 5′-TACCAACAGGTTTGCAAGA(T→A)C-3′ (forward) and 5′-GACTCTACTTTTACCTCGTC-3′ (reverse); and 5′-CTTGCTTTCTGATATAATTTGA(T→A)-3′ (forward) and 5′-GAAGCAGTTCCAACATTTCA-3′ (reverse), respectively. The PCR reactions were done in a total volume of 20 μL that contained 100 ng genomic DNA, 25 pmol/L of each primer, 0.2 mmol/L deoxynucleotide triphosphates, 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, and 1 unit of Taq polymerase (Takara Shuzo Co., Otsu, Shiga, Japan). The PCR cycle conditions consisted of an initial denaturation step at 95°C for 5 minutes followed by 35 cycles of 20 seconds at 94°C; 20 seconds at 58°C for -118T>C, 54°C for IVS1+9G>C, 57°C for IVS10+12A>G, and 49°C for IVS12-6T>C; 20 seconds at 72°C; and a final elongation step at 72°C for 10 minutes. The PCR products were digested overnight at 65°C (−118T>C) or 37°C (IVS1+9G>C, IVS10+12A>G, and IVS12-6T>C) with the appropriate restriction enzymes (New England BioLabs, Beverly, MA). The restriction enzymes for the −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12−6T>C genotypes were *Tsp509*I, *Bgl*I, *HpyCH4*IV, and *Dpn*II, respectively. The digested PCR products were resolved on 6% acrylamide gels

Table 1. Known polymorphisms in the *hMSH2* gene

Location	Genomic	Nucleotide/amino acid change	Frequency of variant allele in previous studies
Promoter	1512	−118T>C	0.20*
Exon 6	14756	G>A/Gly322Asp	0.01†,‡
Exon 11	69413	T>C/Leu556Leu	0.01†
Exon 11	69484	A>G/Lys575Lys	0.05‡
Exon 12	73499	A>G/Asn596Ser	0.00§
Exon 16	81369	T>C/Phe922Phe	0.00‖
Intron 1	1849	IVS1+9G>C	0.62¶, 0.47**
Intron 6	28170	IVS6−10T>C	0.05††
Intron 9	65092	IVS9−9T>A	ND,‡‡ 0.20††
Intron 10	65263	IVS10+12A>G	0.54,‡ 0.35,†† 0.47§§
Intron 12	74808	IVS12−6T>C	0.14,†† 0.11,‖‖ 0.05,¶¶ 0.08,*** 0.29†††

*Ref. 26; 84 healthy Japanese.
†Ref. 22; 46 German colorectal cancer patients.
‡Ref. 29; 100 healthy Spanish.
§Ref. 24; found in Italian hereditary nonpolyposis colorectal cancer family members, but not found in 113 healthy Italians.
‖Ref. 31; found in Portuguese HPNCC family members but not found in 50 healthy Portuguese.
¶Ref. 21: 106 healthy individuals resided in United Kingdom.
**Ref. 23; 57 healthy individuals resided in United States.
††Ref. 20; 30 healthy Norwegians.
‡‡Ref. 30; not determined.
§§Ref. 25; 61 Spanish hereditary nonpolyposis colorectal cancer family members and 50 healthy Spanish.
‖‖Ref. 27; 75 healthy Spanish.
¶¶Ref. 28; 50 healthy Ecuadorians.
***Ref. 32; 837 healthy Finnish.
†††Ref. 33; 487 cancer-free Japanese.

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

and they were stained with ethidium bromide for visualization under UV light. To ensure quality control, the genotyping analysis was done blind with respect to the case/control status. About 10% of the samples were randomly selected to be genotyped again by a different author and the results were 100% concordant. To confirm the genotyping results, selected PCR-amplified DNA samples (n = 2, respectively, for each genotype) were examined by DNA sequencing, and the results were also 100% concordant.

Statistical Analysis. The cases and controls were compared by using Student's t test for the continuous variables and by using the χ^2 test for the categorical variables. Hardy-Weinberg equilibrium was tested for with a goodness-of-fit χ^2 test with 1 degree of freedom to compare the observed genotype frequencies among the subjects with the expected genotype frequencies. The linkage disequilibriums (LD) among the polymorphisms were examined using the Lewontin's standardized coefficient D' ($|D'|$; ref. 36). The haplotypes and their frequencies were estimated based on a Bayesian algorithm using the Phase program (37).[6] Unconditional logistic regression analysis was used to calculate the odds ratios (OR) and 95% confidence intervals (95% CI), with adjustment being made for the possible confounders (gender and family history of lung cancer as a nominal variable; age and pack-years, as continuous variables). In addition to the overall association analysis, we did stratified analyses according to age, gender, smoking status, and tumor histology to further explore the association between the *hMSH2* genotypes/haplotypes and the risk of lung cancer in each stratum. To test which polymorphism is more likely to be the main cause of the observed association, we employed logistic regression models in which we allowed for the effects of the four polymorphisms, both individually and jointly. When multiple comparisons were made, the corrected P values (P_c values) were also calculated for multiple testing by using Bonferroni's inequality method. All the analyses were done using Statistical Analysis Software for Windows, version 8.12 (SAS Institute, Gary, NC).

Results

The demographics of the cases and controls enrolled in this study are shown in Table 2. There were no significant differences between the cases and controls for the mean age or gender distribution, suggesting that the matching based on these two variables was adequate. Cases had a higher prevalence of current smokers than did the controls (P < 0.001) and the number of pack-years in the smokers was significantly higher in the cases than in the controls (39.9 ± 17.9 versus 34.4 ± 17.6 pack-years, respectively; P < 0.001). These differences were controlled for later by the multivariate analyses.

The genotype and polymorphic allele frequencies of the four *hMSH2* polymorphisms (−118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12−6T>C) among the controls and cases are shown in Table 3. The genotype distributions of the four polymorphisms among the controls were in Hardy-Weinberg equilibrium. When the overall lung cancer cases were compared with the controls, no significant difference was found in the distributions of the genotypes for any of the polymorphisms studied. However, when the lung cancer cases were categorized by tumor histology, the distribution of the IVS10+12A>G genotypes in the adenocarcinoma cases differed from that of the controls (P = 0.03), and the variant allele frequencies of the IVS10+12A>G and IVS12-6T>C polymorphisms in the adenocarcinomas differed significantly from those of the controls, respectively (0.305 versus 0.392, P = 0.008; and 0.369 versus 0.304, P = 0.044, respectively).

Table 2. Characteristics of the study population

Variable	Cases (n = 432)	Controls (n = 432)	P
Age (y)	61.6 ± 9.0	60.9 ± 9.3	0.724*
Gender			
Male	352 (81.5)†	352 (81.5)	1.000‡
Female	80 (18.5)	80 (18.5)	
Smoking status			
Current	317 (73.4)	229 (53.0)	0.001‡
Former	39 (9.0)	98 (22.7)	
Never	76 (17.6)	105 (24.3)	
Pack-years§	39.9 ± 17.9	34.4 ± 17.6	<0.001*

*t test.
†Numbers in parentheses, column percentage.
‡χ^2 test.
§In current and former smokers.

Table 4 shows the lung cancer risk related to the *hMSH2* −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12−6T>C genotypes, respectively. No significant association was found between all the four polymorphisms and the risk of overall lung cancer. However, when the lung cancer cases were categorized by tumor histology, those individuals with at least one IVS10+12G allele were at a significantly decreased risk of adenocarcinoma compared with the IVS10+12A homozygotes (adjusted OR, 0.59; 95% CI, 0.40-0.88; P = 0.01), and those individuals with at least one IVS12−6C allele were at a significantly increased risk of adenocarcinoma compared with the IVS12−6T homozygotes (adjusted OR, 1.52; 95% CI, 1.02-2.27; P = 0.04).

The association between the combined genotypes of the -1263A>G and -712C>T polymorphisms and the risk of lung cancer was examined because these two polymorphisms were associated with the risk of lung cancer in a logistic regression analysis for each polymorphism. The distribution of the combined genotypes among the adenocarcinoma cases was significantly different from that among the controls (Table 5; P = 0.02). The IVS10+12 AG+GG/IVS12-6 TT genotype was associated with a significantly lower risk of adenocarcinoma compared with the IVS10+12 AA/IVS12-6 TC+CC genotype (adjusted OR, 0.47; 95% CI, 0.28-0.78; P = 0.004 and P_c = 0.012). In addition, the risk of adenocarcinoma was decreased with decreasing number of high-risk genotypes (P_{trend} = 0.007).

The four *hMSH2* polymorphisms were in LD ($|D'|$ value ranged from 0.60 to 0.83), and we observed 14 haplotypes out of the possible 16 (2^4) haplotypes. For statistical advantage, nine haplotypes that had a frequency of <2% were excluded from any further analysis. The remaining five haplotypes accounted for 92.8% of the chromosomes for the 864 subjects (controls, 93.8%; cases, 91.8%). Table 6 shows the inferred haplotype distributions for the controls and cases, as well as the lung cancer risk as related to the haplotypes. Consistent with genotyping analysis, the TGGT haplotype was associated with a significantly decreased risk of adenocarcinoma compared with the TGAC haplotype (adjusted OR, 0.49; 95% CI, 0.30-0.78; P = 0.003 and P_c = 0.012).

The risk of adenocarcinoma related to the *hMSH2* haplotypes was further examined with stratification according to age, gender, and the smoking status. Because the number of adenocarcinoma cases in each stratum was small, the five haplotypes were categorized into three groups according to the presence of two, one, or no risk alleles at the IVS10+12A>G and IVS12−6T>C (i.e., TGAC, TGAT+CGAT, and TGGT+TGCT). As compared with the TGAC haplotype, the combined haplotype (TGGT+TGCT) with no risk allele was associated with a significantly decreased risk of adenocarcinoma in only

[6]Available at http://www.stat.washington.edu/stephens/software.html

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

Table 3. Genotype frequencies of *hMSH2* polymorphisms in lung cancer cases and controls

Polymorphism	Variables	Genotype*			Polymorphic allele frequency
		1/1	1/2	2/2	
−118T>C	Controls	295 (68.3)	118 (27.3)	19 (4.4)	0.181
	All cases	298 (69.0)	117 (27.1)	17 (3.9)	0.175
	Squamous cell carcinoma	144 (68.6)	54 (25.7)	12 (5.7)	0.186
	Adenocarcinoma	94 (66.7)	43 (30.5)	4 (2.8)	0.181
	Large-cell cancer	6 (75.0)	1 (12.5)	1 (12.5)	0.187
	Small-cell cancer	54 (74.0)	19 (26.0)	0 (0.0)	0.130
IVS1+9G>C	Controls	277 (64.1)	137 (31.7)	18 (4.2)	0.200
	All cases	255 (59.0)	150 (34.7)	27 (6.3)	0.236
	Squamous cell carcinoma	124 (59.1)	70 (33.3)	16 (7.6)	0.243
	Adenocarcinoma	89 (63.1)	45 (31.9)	7 (5.0)	0.209
	Large-cell cancer	3 (37.5)	5 (62.5)	0 (0.0)	0.313
	Small-cell cancer	39 (53.4)	30 (41.1)	4 (5.5)	0.260
IVS10+12A>G	Controls	157 (36.3)	211 (48.8)	64 (14.8)	0.392
	All cases	167 (38.7)	210 (48.6)	55 (12.7)	0.370
	Squamous cell carcinoma	72 (34.3)	108 (51.4)	30 (14.3)	0.400
	Adenocarcinoma	68 (48.2)	60 (42.6)	13 (9.2)†	0.305‡
	Large-cell cancer	3 (37.5)	3 (37.5)	2 (25.0)	0.437
	Small-cell cancer	24 (32.9)	39 (53.4)	10 (13.7)	0.404
IVS12−6T>C	Controls	212 (49.1)	177 (41.0)	43 (10.0)	0.304
	All cases	206 (47.7)	185 (42.8)	41 (9.5)	0.308
	Squamous cell carcinoma	109 (51.9)	83 (39.5)	18 (8.6)	0.283
	Adenocarcinoma	55 (39.1)	68 (48.2)	18 (12.8)	0.369§
	Large-cell cancer	4 (50.0)	4 (50.0)	0 (0.0)	0.250
	Small-cell cancer	38 (52.1)	30 (41.1)	5 (6.9)	0.274

*Wild-type allele is denoted by 1 and the polymorphic allele by 2.
†$P = 0.03$, difference from controls.
‡$P = 0.008$, difference from controls.
§$P = 0.044$, difference from controls.

the younger individuals and the never smokers (adjusted OR, 0.45; 95% CI, 0.27-0.75; $P = 0.002$ and $P_c = 0.004$; and adjusted OR, 0.44; 95% CI, 0.23-0.85; $P = 0.014$ and $P_c = 0.028$, respectively), whereas it had no significant effect on the risk of adenocarcinoma in the older individuals and the ever-smokers (adjusted OR, 0.86; 95% CI, 0.51-1.44; and adjusted OR, 0.74; 95% CI, 0.49-1.12).

Discussion

DNA sequence variations in the *hMSH2* gene may have an effect on the DNA repair capacity, thereby causing inter-individual differences in the susceptibility to lung cancer. To test this hypothesis, we evaluated the potential association of the *hMSH2* polymorphisms (−118T>C, IVS1+9G>C,

Table 4. Adjusted ORs (95% CIs) for lung cancer associated *hMSH2* genotype

Polymorphism	Variables*	Genotype†		
		1/1	1/2	2/2
−118T>C	All cases	1.0	1.03 (0.75-1.40)	0.88 (0.44-1.75)
	Squamous cell carcinoma	1.0	1.01 (0.68-1.50)	1.17 (0.54-2.54)
	Adenocarcinoma	1.0	1.14 (0.74-1.75)	0.80 (0.26-2.45)
	Small-cell cancer	1.0	0.90 (0.51-1.61)	—
IVS1+9G>C	All cases	1.0	1.12 (0.84-1.50)	1.63 (0.87-3.06)
	Squamous cell carcinoma	1.0	1.05 (0.72-1.52)	2.16 (1.04-4.47)‡
	Adenocarcinoma	1.0	1.01 (0.66-1.55)	1.12 (0.44-2.83)
	Small-cell cancer	1.0	1.47 (0.87-2.50)	1.71 (0.54-5.47)
IVS10+12A>G	All cases	1.0	0.93 (0.69-1.25)	0.81 (0.53-1.24)
	Squamous cell carcinoma	1.0	1.18 (0.81-1.72)	1.09 (0.64-1.85)
	Adenocarcinoma	1.0	0.64 (0.42-0.96)§	0.45 (0.23-0.88)‖
		1.0	0.59 (0.40-0.88)¶,**	
	Small-cell cancer	1.0	1.25 (0.72-2.19)	1.06 (0.47-2.38)
IVS12−6T>C	All cases	1.0	1.10 (0.82-1.46)	0.97 (0.60-1.57)
	Squamous cell carcinoma	1.0	0.90 (0.63-1.30)	0.77 (0.41-1.45)
	Adenocarcinoma	1.0	1.51 (1.00-2.30)	1.56 (0.82-2.98)
		1.0	1.52 (1.02-2.27)‡,¶	
	Small-cell cancer	1.0	0.98 (0.58-1.66)	0.63 (0.23-1.74)

NOTE: ORs (95% CIs) were adjusted for age, gender, and pack-years of smoking.
*The number of cases in each stratum: all, 432: squamous cell carcinoma, 210; adenocarcinoma, 141; and small-cell cancer, 73. The number of cases in each genotype is same as that in Table 3.
†Wild-type allele is denoted by l and the polymorphic allele by 2.
‡$P = 0.04$.
§$P = 0.03$.
‖$P = 0.02$.
¶Dominant model for the variant allele (1/2 + 2/2 versus 1/1).
**$P = 0.01$.

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

Table 5. Combined genotype frequencies of *hMSH2* IVS10+12A>G and IVS12−6T>C polymorphisms among the cases and controls and their associations with the risk of lung cancer

Combined genotype		Controls	All cases		Adenocarcinoma cases	
		no. (%)	no. (%)	Adjusted* OR (95% CI)	no. (%)	Adjusted* OR (95% CI)
IVS10+12A>G	IVS12−6T>C					
AA	TC+CC	110 (25.5)	115 (26.6)	1.0	48 (34.0)†	1.0
AA	TT	47 (10.9)	52 (12.0)	1.04 (0.64-1.69)	20 (14.2)	1.00 (0.53-1.90)
AG+GG	TC+CC	110 (25.5)	111 (25.7)	0.97 (0.66-1.41)	38 (27.0)	0.78 (0.47-1.30)
AG+GG	TT	165 (38.2)	154 (36.7)	0.88 (0.62-1.25)	35 (24.8)	0.47 (0.28-0.78)‡
No. risk genotype						
2 (AA/TC+CC)		110 (25.5)	115 (26.6)	1.0	48 (34.0)§	1.0‖
1 (AA/TT or AG+GG/TC+CC)		157 (36.3)	163 (37.7)	0.99 (0.70-1.40)	58 (41.1)	0.84 (0.53-1.34)‖
0 (AG+GG/TT)		165 (38.2)	154 (36.7)	0.88 (0.62-1.25)	35 (24.8)	0.47 (0.28-0.78)‖,¶

*Adjusted for age, gender, and pack-years of smoking.
†$P = 0.02$, two-sided χ^2 test for the combined genotype distributions between the controls and adenocarcinoma cases.
‡$P = 0.004$ and $P_c = 0.012$.
§$P = 0.01$, two-sided χ^2 test for the combined genotype distributions between the controls and adenocarcinoma cases.
‖$P_{trend} = 0.007$, test for trend of odds were two sided and based on likelihood ratio tests.
¶$P = 0.004$ and $P_c = 0.008$.

IVS10+12A>G, and IVS12−6T>C) with the risk of lung cancer. The *hMSH2* IVS10+12A>G and IVS12−6T>C polymorphisms were significantly associated with the risk of adenocarcinoma of the lung. This finding suggests that the *hMSH2* IVS10+12A>G and IVS12−6T>C polymorphisms could be used as markers for the genetic susceptibility to adenocarcinoma. Of the three major histologic types of lung cancer, the proportion of adenocarcinoma is increasing worldwide. Thus, identification of the genetic factors that are responsible for the susceptibility to adenocarcinoma is indispensable for establishing novel and efficient ways to prevent this disease.

In the current study, we validated the presence of hMSH2 −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12−6T>C polymorphisms in a Korean population. However, the other seven polymorphisms among the previously reported polymorphisms listed in Table 1 were not detected in the preliminary study that included 27 healthy controls. These samples included 54 chromosomes, which provides at least a 95% confident level to detect alleles with frequencies >5%. Thus, it is very likely that if these polymorphisms exist, they may not play a major role in the genetic susceptibility to lung cancer in the Korean population (38, 39). The frequency of the −118C allele among healthy Koreans was 0.181, which was similar to that of Japanese (0.202; ref. 26). Frequencies of the IVS1+9C and IVS10+12G alleles among healthy Koreans were 0.20 and 0.39, respectively, which were lower than those of healthy Caucasians (0.47-0.62 and 0.35-0.54, respectively; refs. 20, 21, 23, 25, 29). Frequency of IVS12−6C allele among healthy Koreans was 0.304, which was similar to that of healthy Japanese (0.285; ref. 33) but significantly lower than that of healthy Caucasians (0.05-0.14; refs. 20, 27, 28, 32).

In the current study, the *hMSH2* polymorphisms were significantly associated with the risk of adenocarcinoma but they were not associated with the squamous cell carcinoma or small-cell carcinoma. Although the reason for the observed histology-dependent difference in the risk conferred by the *hMSH2* polymorphism is unknown, this difference may be attributable to the differences in the pathways of carcinogenesis among the different histologic types of lung cancer. Various lines of evidence have suggested that the histologic type of lung cancer may be determined by the particular initiating agent to which an individual is exposed (40-42). Therefore, certain polymorphisms could confer a greater susceptibility to a particular histologic type of lung cancer (34, 35, 43). Several studies have shown that *hMSH2* gene was frequently inactivated in lung cancers, and the *hMSH2* inactivation rate was significantly higher in adenocarcinoma than in squamous cell carcinoma (44-46). These observations suggest that the *hMSH2* gene may have a pronounced association with the development of adenocarcinoma, and these observations are comparable with our finding that the *hMSH2* polymorphisms play an important role in determining the genetic susceptibility to adenocarcinoma.

Another interesting finding of the present study is that the effect of the *hMSH2* polymorphisms on the risk of adenocarcinoma was pronounced in the younger individuals and never smokers. Several recent studies suggest that adenocarcinomas arising in never smokers and smokers are caused by different

Table 6. Distribution of *hMSH2* inferred haplotypes in controls and cases

Haplotype*	Controls ($n = 810$)†	All cases ($n = 793$)†		Histologic type of lung cancer‡					
	no. (%)	no. (%)	OR§ (95% CI)	Squamous cell carcinoma ($n = 386$)†		Adenocarcinoma ($n = 261$)†		Small-cell cancer ($n = 133$)†	
				no. (%)	OR§ (95% CI)	no. (%)	OR§ (95% CI)	no. (%)	OR§ (95% CI)
TGAC	235 (29.0)	234 (29.5)	1.0	102 (26.4)	1.0	92 (35.3)	1.0	36 (27.2)	1.0
TGAT	141 (17.4)	143 (18.0)	1.01 (0.74-1.36)	68 (17.6)	1.12 (0.76-1.64)	47 (18.0)	0.86 (0.57-1.31)	27 (20.3)	1.21 (0.70-2.10)
CGAT	128 (15.8)	124 (15.6)	0.96 (0.70-1.31)	63 (16.3)	1.14 (0.77-1.69)	44 (16.9)	0.86 (0.56-1.33)	15 (11.3)	0.75 (0.39-1.44)
TGGT	151 (18.6)	127 (16.0)	0.84 (0.62-1.13)	69 (17.9)	1.11 (0.76-1.62)	30 (11.5)	0.49 (0.30-0.78)‖	24 (18.1)	1.07 (0.61-1.88)
TCGT	155 (19.1)	165 (20.8)	1.04 (0.78-1.39)	84 (21.8)	1.25 (0.86-1.80)	48 (18.4)	0.77 (0.51-1.17)	31 (23.3)	1.28 (0.75-2.18)

*The order of polymorphisms for the haplotypes is as follows: −118T>C, IVS1+9G>C, IVS10+12A>G, and IVS12−6T>C.
†Nine haplotypes that had a frequency of <2% were excluded from analysis; 54 controls and 71 cases (34 squamous cell carcinoma, 21 adenocarcinoma, 3 large-cell cancer, and 13 small-cell cancer), respectively.
‡Eight large-cell carcinoma cases were excluded from analysis.
§Adjusted for age, gender, pack-years of smoking, and family history of lung cancer.
‖$P = 0.003$ and $P_c = 0.012$.

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

etiologies (i.e., carcinogens other than environmental tobacco smoke play an important role in the pathogenesis of adenocarcinomas in never smokers; refs. 47, 48). In view of this suggestion, our finding that the effect of the *hMSH2* polymorphisms on the risk of adenocarcinoma is more pronounced in the never smokers may be due to that the *hMSH2* polymorphisms may contribute to the development of adenocarcinoma by influencing on the repair capacity of DNA damages caused by environmental factor(s) other than cigarette smoking and/or endogenous factor(s). This explanation is comparable with the previous studies (45, 46) that showed that the altered hMSH2 protein expression in non–small-cell lung cancers was significantly associated with the adenocarcinoma histology, younger patients, female patients, and the never smokers. However, it is possible that such a finding is attributable to chance because of the relatively small numbers of subjects in the subgroups. Additional studies with more patients will be needed to confirm this finding.

The haplotypes can increase the power to detect disease associations compared with a single polymorphism on account of the higher heterozygosity and tighter LD with the disease-causative variant (49-51). In this study, the IVS1+9G>C polymorphism had no effect on the risk of adenocarcinoma in individual polymorphism analysis, but haplotype analysis showed that the TGGT haplotype carrying the IVS1+9G allele was associated with a significantly decreased risk compared with the TGAC haplotype whereas the TCGT haplotype carrying IVS1+9C allele was not significantly associated with the risk of adenocarcinoma. These results also suggest that haplotype analysis may be a more suitable tool for assessing the disease association than the individual polymorphism.

Several polymorphisms have been reported in the *hMSH2* gene (20-33). It has been suggested that one of these polymorphism, IVS12−6T>C, may predispose to some kind of cancer (27, 28, 32, 52, 53). In agreement with the previous studies, the IVS12−6C allele was associated with a significantly increased risk of lung adenocarcinoma in the present study. A characteristic finding in the present study is that the IVS10+12A>G polymorphism as well as the IVS12−6T>C polymorphism was significantly associated with the risk of adenocarcinoma. Because these two polymorphisms were in strong LD, it is not easy to discern the relative contribution of each polymorphism to the observed association. In an attempt to resolve this, we compared three different logistic regression models (each polymorphism alone and both polymorphisms together). The model incorporating both polymorphisms did not fit significantly better than the model with the IVS10+12A>G alone ($P = 0.15$) but the model with the IVS12−6T>C alone fitted less well than the model incorporating both polymorphisms ($P = 0.03$). These results suggest that the genetic effect of the IVS10+12A>G polymorphism is stronger than the effect of the IVS12−6T>C polymorphism.

It has been reported that multiple alternatively spliced isoforms of hMSH2 mRNA are present in normal human tissues, some of them would produce truncated proteins that lack exons 2-8 or exon 13 and others were in-frame deletions lacking exon 5 or exons 2-7; the relative expression level of each splicing variant has shown considerable interindividual variation (54-56). Although the functional significance of the IVS10+12A>G and IVS12−6T>C polymorphisms remains to be elucidated, these splicing site polymorphisms may have an influence on the alternative splicing of *hMSH2*, resulting in interindividual variation in the expression levels of the hMSH2 splicing variants. Some of these variant proteins might cause a dominant-negative effect by interfering with the wild-type hMSH2 protein. An alternative explanation for the association between the *hMSH2* polymorphisms and the risk of lung cancer may be due to LD with either another *hMSH2* variant or with an adjacent true susceptible gene.

In conclusion, we found that the *hMSH2* polymorphisms and their haplotypes were associated with the risk of adenocarcinoma of the lung. The effects of the *hMSH2* polymorphisms on the risk of adenocarcinoma were more evident in the younger individuals and the never smokers. These results suggest that the *hMSH2* gene may contribute to an inherited predisposition to adenocarcinoma of the lung. It is possible that these findings, particularly those findings from stratified analyses, can be attributed to chance because of the relatively small numbers of cases in the subgroups. Therefore, additional studies with larger sample sizes are required to confirm our findings. Moreover, because the gene-gene interactions and gene-environment interactions often vary between ethnic groups, further studies are needed to clarify the association between the *hMSH2* polymorphisms and lung cancer in diverse ethnic populations.

References

1. Shields PG, Harris CC. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 2000;18:2309–15.
2. Mohrenweiser HW, Jones IM. Variation in DNA repair is a factor in cancer susceptibility: a paradigm for the promise and perils of individuals and population risk estimation? Mutat Res 1998;400:15–24.
3. Wood EC, Mitchell M, Sgouros J, et al. Human DNA repair genes. Science 2001;291:1284–8.
4. Bernstein C, Bernstein H, Payne CM, et al. DNA repair/pro-apoptotic dual role proteins in five major DNA repair pathways: fail-safe protection against carcinogenesis. Mutat Res 2002;511:145–78.
5. Cheng L, Eicher SA, Guo Z, et al. Reduced DNA repair capacity in head and neck cancer patients. Cancer Epidemiol Biomarkers Prev 1998;7:465–8.
6. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002; 11:1513–30.
7. Park JY, Lee SY, Jeon HS, et al. Polymorphisms of the DNA repair gene *XRCC1* and risk of primary lung cancer. Cancer Epidemiol Biomarkers Prev 2002;11:23–7.
8. Park JY, Lee SY, Jeon H-S, et al. Lys751Gln polymorphism of the DNA repair gene *XPD* and risk of primary lung cancer. Lung Cancer 2002;36: 15–6.
9. Park JY, Park SH, Choi JE, et al. Polymorphism of the DNA repair gene *XPA* and risk of primary lung cancer. Cancer Epidemiol Biomarkers Prev 2002;11:993–7.
10. Kolodner RD. Mismatch repair: mechanisms and relationship to cancer susceptibility. Trends Biochem Sci 1995;20:397–401.
11. Bellacosa A. Functional interactions and signaling properties of mammalian DNA mismatch repair proteins. Cell Death Differ 2004;25:1821–7.
12. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair deficient human cancers: toward a new concept of target genes for instability. Cancer Res 2002;62:2447–54.
13. Simpkins SB, Bocker T, Swisher EM, et al. *MLH1* promoter methylation and gene silencing is the primary cause of microsatellite instability in sporadic cancers. Hum Mutat Genet 1999;8:661–6.
14. Boland CR, Thibodeau SN, Hamilton SR, et al. A national cancer institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998; 58:5248–57.
15. Park JY, Jeon H-S, Park SH, et al. Microsatellite alteration in histologically normal lung tissue of patients with non-small cell lung cancer. Lung Cancer 2000;30:83–9.
16. Wu J, Gu L, Wang H, et al. Mismatch repair processing of carcinogen-DNA adducts triggers apoptosis. Mol Cell Biol 1999;19:8292–301.
17. Hickman MJ, Samson LD. Role of DNA mismatch repair and p53 in signaling induction of apoptosis by alkylating agents. Proc Natl Acad Sci 1999;96:10764–9.
18. Mellon I, Rajpal DK, Koi M, et al. Transription-coupled deficiency and mutations in mismatch repair genes. Science 1996;272:557–60.
19. Bertrand P, Tishkoff DX, Filosi N, et al. Physical interaction between components of DNA mismatch repair and nucleotide excision repair. Proc Natl Acad Sci 1998;95:14278–83.
20. Borresen AL, Lothe RA, Meling GI, et al. Somatic mutations in the *hMSH2* gene in microsatellite unstable colorectal carcinomas. Hum Mol Genet 1995;4:2065–72.
21. Bubb VJ, Curtis LJ, Cunningham C, et al. Microsatellite instability and the role of hMSH2 in sporadic colorectal cancer. Oncogene 1996;12:2641–9.
22. Wehner M, Buschhausen L, Lamberti C, et al. Hereditary nonpolyposis colorectal cancer (HNPCC): eight germline mutations in *hMSH2* or *hMLH1* genes. Hum Mutat 1997;10:241–4.
23. Herfarth KKF, Kodner IJ, Whelan AJ, et al. Mutations in *MLH1* are more frequent than in *MSH2* in sporadic colorectal cancers with microsatellite instability. Genes Chromosomes Cancer 1997;18:42–9.

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

24. Viel A, Genuardi M, Capozzi E, et al. Characterization of *MSH2* and *MLH1* mutations in Italian families with hereditary nonpolyposis colorectal cancer. Genes Chromosomes Cancer 1997;18:8–18.
25. Wahlberg SS, Nystrom-Lahti M, Kane MF, et al. Low frequency of hMSH2 mutations in Swedish HNPCC families. Int J Cancer 1997;74:134–7.
26. Iwahashi Y, Ito E, Yanagisawa Y, et al. Promoter analysis of the human mismatch repair gene *hMSH2*. Gene 1998;213:141–7.
27. Palicio M, Blanco I, Tortola S, et al. Intron splice acceptor site polymorphism in the *hMSH2* gene in sporadic and familial colorectal cancer. Br J Cancer 2000;82:535–7.
28. Paz-y-Mino C, Perez JC, Fiallo BF, et al. A polymorphism in the *hMSH2* gene (gIVS12−6T>C) associated with non-Hodgkin lymphomas. Cancer Genet Cytogenet 2002;133:29–33.
29. Caldes T, Godino J, de la Hoya M, et al. Prevalence of germline mutations in *MLH1* and *MSH2* in hereditary nonpolyposis colorectal cancer families from Spain. Int J Cancer 2002;98:774–9.
30. Kurzawski G, Safranow K, Suchy J, et al. Mutation analysis of *MLH1* and *MSH2* genes performed by denaturing high-performance liquid chromatography. J Biochem Biophys Methods 2002;51:89–100.
31. Isidro G, Matos S, Goncalves V, et al. Novel MLH1 mutations and a novel MSH2 polymorphism identified by SSCP and DHPLC in Portuguese HNPCC families. Hum Mutat 2003;22:419–20.
32. Worrillow LJ, Travis LB, Smith AG, et al. An intron splice acceptor polymorphism in *hMSH2* and risk of leukemia after treatment with chemotherapeutic alkylating agents. Clin Cancer Res 2003;9:3012–20.
33. Hishida A, Matsuo K, Hamajima N, et al. Polymorphism in the hMSH2 gene (gIVS12−6T→C) and risk of non-Hodgkin lymphoma in a Japanese population. Cancer Genet Cytogenet 2003;147:71–4.
34. Lee SJ, Jeon H-S, Jang JS, et al. *DNMT3B* polymorphisms and risk of primary lung cancer. Carcinogenesis 2005;26:403–9.
35. Lee SJ, Lee SY, Jeon H-S, et al. *Vascular endothelial growth factor* gene polymorphisms and risk of primary lung cancer. Cancer Epidemiol Biomarkers Prev 2005;14:571–5.
36. Lewontin RC. On measures of gametic disequilibrium. Genetics 1998;120: 849–52.
37. Stephens M, Smith MJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001;68:978–89.
38. Risch NJ. Searching for genetic determinants in the new millennium. Nature 2000;405:847–56.
39. Rebbeck TR, Ambrosone CB, Bell DA, et al. SNPs, haplotypes, and cancer: applications in molecular epidemiology. Cancer Epidemiol Biomarkers Prev 2004;13:681–7.
40. Deutsch-Wenzel R, Brune H, Grimmer G, et al. Experimental studies in rat lungs on the carcinogenicity and dose-response relationships of eight frequently occurring environmental polyaromatic hydrocarbons. J Natl Cancer Inst 1983;71:539–44.
41. Hoffman D, Rivenson A, Murphy SE, et al. Cigarette smoking and adenocarcinoma of the lung: the relevance of nicotine-derived nitrosamines. J Smoking Relat Disord 1993;4:165–90.
42. Smith CJ, Livingston SD, Doolittle DJ. An international literature survey of ''IARC group I carcinogens'' reported mainstream cigarette smoke. Food Chem Toxicol 1997;35:1107–30.
43. Gu J, Spitz MR, Yang F, et al. Ethnic differences in poly(ADP-ribose) polymerase pseudogene genotype distribution and association with lung cancer risk. Carcinogenesis 1999;20:1465–9.
44. Xinarianos G, Liloglou T, Prime W, et al. hMLH1 and hMSH2 expression correlates with allelic imbalance on chromosome 3p in non-small cell lung cancer. Cancer Res 2000;60:4216–21.
45. Wang YC, Lu YP, Tseng RC, et al. Inactivation of *hMLH1* and *hMSH2* by promoter methylation in primary non-small cell lung tumors and matched sputum samples. J Clin Invest 2003;111:887–95.
46. Hsu HS, Wen CK, Tang YA, et al. Promoter hypermethylation is the predominant mechanism in *hMLH1* and *hMSH2* deregulation and is a poor prognostic factor in nonsmoking lung cancer. Clin Cancer Res 2005;11: 5410–6.
47. Marchetti A, Martella C, Felicioni L, et al. *EGFR* mutations in non-small cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. J Clin Oncol 2005;23:857–65.
48. Gazdar AF, Shigematsu H, Herz J, et al. Mutations and addiction to *EGFR*: the Achilles ''heal'' of lung cancers? Trends Mol Med 2004;10:481–6.
49. Judson R, Stephens JC, Windemuth A. The predictive power of haplotypes in clinical response. Pharmacogenomics 2000;1:15–26.
50. Judson R, Stephens JC. Notes from the SNP vs. haplotype front. Pharmacogenomics 2001;2:7–10.
51. Leng S, Cheng J, Zhang L, et al. The association of XRCC1 haplotypes and chromosomal damage levels in peripheral blood lymphocytes among coke-oven workers. Cancer Epidemiol Biomarkers Prev 2005;14:1295–301.
52. Brentnall TA, Rubin CE, Crispin DA, et al. A germline substitution in the human *MSH2* gene is associated with high-grade dysplasia and cancer in ulcerative colitis. Gastroenterology 1995;109:151–5.
53. Goessl C, Plaschke J, Pistorius S, et al. An intronic germline transition in the HPNCC gene *hMSH2* is associated with sporadic colorectal cancer. Eur J Cancer 1997;30A:1550–2.
54. Xia L, Shen W, Ritacca F, et al. A truncated *hMSH2* transcript occurs as a common variant in the population: implication for genetic diagnosis. Cancer Res 1996;56:2289–92.
55. Mori Y, Shiwaku H, Fukushige S, et al. Alternative splicing of *hMSH2* in normal human tissues. Hum Genet 1997;99:590–5.
56. Genuardi M, Viel A, Bonora D, et al. Characterization of *MLH1* and *MSH2* alternative splicing and its relevance to molecular testing of colorectal cancer. Hum Genet 1998;102:15–20.

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Polymorphisms in the *hMSH2* Gene and the Risk of Primary Lung Cancer

Chi Young Jung, Jin Eun Choi, Jung Min Park, et al.

Cancer Epidemiol Biomarkers Prev 2006;15:762-768.

Updated version	Access the most recent version of this article at: http://cebp.aacrjournals.org/content/15/4/762

Cited articles	This article cites 56 articles, 22 of which you can access for free at: http://cebp.aacrjournals.org/content/15/4/762.full.html#ref-list-1
Citing articles	This article has been cited by 6 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/15/4/762.full.html#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.

Downloaded from cebp.aacrjournals.org on February 2, 2016. © 2006 American Association for Cancer Research.